

CLAIM AMENDMENTS

1. (Currently Amended) A computer-implemented computer implemented system that facilitates analyzing a degree of relatedness between newsgroups newgroup clusters, comprising the following computer executable components:

one or more processors;

memory coupled to the one or more processors;

a data reception component, stored in the memory and executed by the one or more processors, that receives and recognizes data relating to a plurality of the newsgroups and cross-postings between the newsgroups; and

[[an]] a graphing engine, stored in the memory and executed by the one or more processors, that constructs a weighted graph with a subset of the newsgroups represented as vertices of the weighted graph[[.]] and cross-postings between two of relating to the subset of newsgroups represented as an edge between vertices corresponding to the two newsgroups edges.

2. (Cancelled)

3. (Currently Amended) The system of claim 1, further comprising a segmenting component, stored in the memory and executed by the one or more processors, that segments the weighted graph via spectral clustering.

4. (Currently Amended) The system of claim 3, wherein the segmenting is performed at least partially as a function of a number of cross-postings between newsgroups.

5. (Currently Amended) The system of claim 4, wherein the segmenting component partitions partitioning vertices of the weighted graph into segments so that a total number of edges between different segments is substantially minimized.

6. (Original) The system of claim 5, wherein the segmenting component partitions segments recursively.

7. (Currently Amended) The system of claim 3, further comprising a post-processing-component, stored in the memory and executed by the one or more processors, that merges a first cluster of vertices and edges of the weighted graph into a second cluster of vertices and edges of the weighted graph if a sum of weights between the clusters is greater than a threshold.

8. (Currently Amended) The system of claim 7, wherein the threshold being is at least partially a function of a sum of weights of an edge edges adjacent to the first cluster.

9. (Currently Amended) The system of claim 8, wherein a first cluster and a second cluster two-clusters are merged when a sum of the weights of edges

between [[a]] the first cluster and [[a]] the second cluster is more than half of a sum of weights of edges adjacent to the first cluster.

10. (Currently Amended) The system of claim 1, further comprising a filtering component, stored in the memory and executed by the one or more processors, that facilitates excluding particular newsgroups from being represented in the weighted graph so as to facilitate reducing [[the]] a size of the graph.

11. (Currently Amended) The system of claim 10, wherein the filtering component excludes newsgroups which do not contain fewer than a threshold number of postings.

12. (Currently Amended) The system of claim 10, wherein the filtering component excludes newsgroups by utilizing an implicitly trained classifier that infers the type of newsgroup desired by a user based at least in part on newsgroups previously accessed by the user.

13. (Currently Amended) The system of claim 1, further comprising a paring component, stored in the memory and executed by the one or more processors, that removes trims edges of the weighted graph with a weight less than a threshold weight.

14. (Currently Amended) The system of claim 13, wherein the threshold weight is an increasing function of a number of the newsgroups size of the data to be graphed.

15. (Currently Amended) The system of claim 14, wherein the paring component removes vertices when the vertices are [[not]] interconnected to other vertices by less than edges to a threshold number of edges vertices.

16-17. (Canceled)

18. (Currently Amended) They The system of claim 1, further comprising outputs the weighted graph to a display device configured to display an output of the weighted graph.

19. (Currently Amended) The system of claim 18, wherein the output of displays the weighted graph comprises a textual output textually.

20. (Canceled)

21. (Currently Amended) A computer implemented method for creating a weighted newsgroup graph, the method comprising the following computer executable acts:

receiving and recognizing data relating to a plurality of newsgroups and cross-posts between the newsgroups; and

constructing, by one or more processors, a weighted graph such that newsgroups are represented as vertices and cross-posts between two of the newsgroups are represented as edges between two of the vertices corresponding to the two newsgroups.

22. (Currently Amended) The method of claim 21, further comprising excluding one or more newsgroups from the weighted graph when the one or more newsgroups does not contain fewer than a threshold number of postings.

23. (Currently Amended) The method of claim 21, further comprising excluding one or more newsgroups from the weighted graph by utilizing implicitly trained classifiers that infer the type of newsgroups desired by a user based at least in part on newsgroups previously accessed by the user.

24. (Original) The method of claim 21, further comprising segmenting the weighted graph into clusters.

25. (Original) The method of claim 24, wherein a spectral clustering algorithm is utilized to segment the weighted graph into clusters.

26. (Original) The method of claim 25, wherein the spectral clustering algorithm is applied recursively to the weighted graph.

27. (Original) The method of claim 26, wherein the spectral clustering algorithm comprises:

calculating vector v by solving an equation $Lv = \lambda Dv$, wherein $L = D - A$ is the Laplacian of the adjacency matrix $A = (a_{ij})$, D is a diagonal matrix with $d_{ii} = \sum_j a_{ij}$, and λ is the second smallest eigenvalue of L ;

determining maximum and minimum values contained within vector v ;

dividing an interval between the maximum and minimum values of v into Q smaller intervals;

locating a smallest $Mcut$ ratio at endpoints of the Q intervals, wherein S and \bar{S} are two segments resulting from a proposed cut, $cut = \sum_{i \in S} a_{ij}, W_s = \sum_{i,j \in S} a_{ij}$, and $Mcut = \frac{cut}{W_s} + \frac{cut}{W_{\bar{S}}}$

calculating a minimum $Mcut$ ratio of an integer P eigenvector entries before and after the endpoint found to have a lowest $Mcut$ ratio of the Q intervals;

comparing the minimum $Mcut$ ratio of the P eigenvector entries to a threshold t ; and

segmenting the eigenvector entry where the minimum $Mcut$ ratio is found if the $Mcut$ ratio is less than the threshold t .

28. (Original) The method of claim 24, further comprising merging the segmented clusters if the weights of edges between clusters is greater than a threshold.

29. (Currently Amended) The method of claim 28, wherein the threshold being is a function of a sum of weights of an edge edges adjacent to the first cluster.

30. (Currently Amended) A computer implemented system that facilitates analyzing newsgroup similarity-clusters, comprising the following computer executable components:

one or more processors;

memory coupled to the one or more processors;

a data reception component, stored in the memory and executed by the one or more processors, that receives data relating to a plurality of newsgroups and cross-postings between the plurality of newsgroups;

[[an]] a graphing engine, stored in the memory and executed by the one or more processors, that constructs a weighted graph with a subset of the newsgroups represented as vertices of the graph[[,]] and cross-postings between two newsgroups of relating to the subset of newsgroups represented as edges between vertices corresponding to the two newsgroups; and further comprising at least one of the following components:

a filtering component, stored in the memory and executed by the one or more processors, that facilitates excluding particular newsgroups from being represented in the weighted graph so as to facilitate reducing [[the]] a size of the weighted graph;

a paring component, stored in the memory and executed by the one or more processors, that trims removes edges of the graph with a weight less than a threshold weight so as to facilitate reducing the size of the graph;

a segmenting component, stored in the memory and executed by the one or more processors, that segments the weighted graph via spectral clustering; and

a post-processing component, stored in the memory and executed by the one or more processors, that merges a first cluster of vertices and edges of the weighted graph into a segment second cluster of vertices and edges of the weighted graph if a sum of weights between the clusters is greater than a threshold.

31. (Currently Amended) The system of claim 30, further comprising a data store for storing ~~at least one of the following:~~

~~news-group~~ newsgroup data received by the data reception component;

algorithms utilized for segmenting the weighted graph;

the weighted graph generated by the graphing engine; and

~~the segmented a-graph upon generated by segmentation of~~ the weighted graph being segmented via the segmenting component.

32. (Currently Amended) The system of claim 30, wherein the post-processing component outputting the outputs a modified weighted graph.

33-39. (Canceled)

40. (Currently Amended) A computer-implemented computer implemented method for creating a cluster graph comprising the following computer executable steps:

receiving newsgroup data including cross-posting data;

excluding newsgroups that ~~do not~~ contain fewer than a threshold number of postings;

~~paring removing edges with a weight below a threshold weight~~;

generating a weighted graph with the newsgroups represented as vertices and the cross-postings between newsgroups represented as edges between the vertices, wherein a weight of the edges is based at least in part on a number of cross-postings between the two newsgroups;

segmenting the weighed graph into clusters; and
merging clusters if the sum of the weights between clusters is greater than a threshold; and
outputting the graph.

41. (Currently Amended) A computer implemented system that facilitates analyzing newsgroup clusters, comprising:

one or more processors; and
memory coupled to the one or more processors, the memory comprising
instructions that cause the one or more processors to perform acts comprising:
means-a data reception component, stored in the memory and executable by the
one for more processors, for receiving and recognizing data relating to a plurality of
newsgroups; and

means-a graphing engine, stored in the memory and executable by the one for
more processors, for constructing a weighted graph with a subset of the newsgroups
represented as vertices of the graph[[],]] and cross-postings relating to the subset of
newsgroups represented as edges, wherein the vertices are weighted based at least in
part on a number of postings to the corresponding newsgroups and the edges are
weighted based at least in part on a number of cross-postings between two
newsgroups.

42. (Cancelled)

43. (New) The system of claim 1, wherein the vertices are weighted based at least in part on a number of postings within the newsgroups.

44. (New) The system of claim 1, wherein the edges are weighted based at least in part on a number of cross-posts between the respective two newsgroups.

45. (New) The system of claim 1, wherein the newsgroups comprise Usenet groups.

46. (New) The system of claim 1, further comprising a search engine configured to use the weighted graph when executing a newsgroup search and providing results from the newsgroups search.

47. (New) The system of claim 1, further comprising an e-mail program configured to generate a suggestion that a post be cross-posted to other newsgroups, the other newsgroups identified at least in part by the weighted graph.

48. (New) the method of claim 21, wherein the vertices are weighted based at least in part on a number of posts to the corresponding newsgroups.

49. (New) The method of claim 21, wherein the edges are weighted based at least in part on cross-posts between two newsgroups.

50. (New) The system of claim 30, wherein the vertices of the weighted graph are weighted based at least in part on a number of postings to the corresponding newsgroups and the edges of the weighted graph are weighted based at least in part on a number of cross-postings between the two corresponding newsgroups.

51. (New) The system of claim 30, further comprising a search engine configured to use the weighted graph when executing a newsgroup search and providing results from the newsgroups search.

52. (New) The system of claim 30, further comprising an e-mail program configured to generate a suggestion that a post be cross-posted to other newsgroups, the other newsgroups identified at least in part by the weighted graph.

53. (New) The method of claim 40, wherein the weight of the vertices is based at least in part on a number of postings within the newsgroups corresponding to the vertices.

54. (New) The method of claim 40, further comprising outputting the graph.